Monte Carlo Studies on Equilibrium Globular Protein Folding. I. Homopolymeric Lattice Models of β-Barrel Proteins

ANDRZEJ KOLINSKI,* JEFFREY SKOLNICK,† and ROBERT YARIS, *Institute of Macromolecular Chemistry, Department of Chemistry, Washington University, St. Louis, Missouri 63130*

Synopsis

Dynamic Monte Carlo studies have been performed on various diamond lattice models of β-proteins. Unlike previous work, no bias toward the native state is introduced; instead, the protein is allowed to freely hunt through all of phase space to find the equilibrium conformation. Thus, these systems may aid in the elucidation of the rules governing protein folding from a given primary sequence; in particular, the interplay of short- vs long-range interaction can be explored. Three distinct models (A–C) were examined. In model A, in addition to the preference for *trans* (t) over *gauche* states (g^+ and g^-) (thereby perhaps favoring β-sheet formation), attractive interactions are allowed between all nonbonded, nearest neighbor pairs of segments. If the molecules possess a relatively large fraction of t states in the denatured form, on cooling spontaneous collapse to a well-defined β-barrel is observed. Unfortunately, in model A the denatured state exhibits too much secondary structure to correctly model the globular protein collapse transition. Thus in models B and C, the local stiffness is reduced. In model B, in the absence of long-range interactions, t and g states are equally weighted, and cooperativity is introduced by favoring formation of adjacent pairs of nonbonded (but not necessarily parallel) t states. While the denatured state of these systems behaves like a random coil, their native globular structure is poorly defined. Model C retains the cooperativity of model B but allows for a slight preference of t over g states in the short-range interactions. Here, the denatured state is indistinguishable from a random coil, and the globular state is a well-defined β-barrel. Over a range of chain lengths, the collapse is well represented by an all-or-none model. Hence, model C possesses the essential qualitative features observed in real globular proteins. These studies strongly suggest that the uniqueness of the globular conformation requires some residual secondary structure to be present in the denatured state.

INTRODUCTION

The equilibrium transition from the denatured to the native state of a protein has sometimes been compared to the collapse transition of synthetic polymers.[1] In the latter case, when the temperature of an extremely dilute polymer solution is lowered, the polymer chains undergo a transition from a voluminous random coil to a dense globular state having a random distribution of segments. For linear flexible macromolecules of finite degree of polymerization, the transition is smooth and is poorly represented by an all-or-none model.[2-6] However, there is theoretical[7-11] and experimental[12] evidence that increasing the stiffness of the chain backbone leads to a substantially steeper

*Permanent address: Department of Chemistry, University of Warsaw, 02-093 Warsaw, Poland.
†Alfred P. Sloan Foundation Fellow.

Biopolymers, Vol. 26, 937–962 (1987)

CCC 0006-3525/87/060937-26$04.00

transition. Compared to the case of synthetic polymers, the conformational transition in proteins is much more complicated. The low-temperature "native" state has a unique structure comprised of stiff pieces of α-helices and/or β-sheets joined by more flexible bends whose tertiary structure is determined by the sequence of amino acids in the polypeptide chain; for the smaller proteins the transition has an *all-or-none* character.[13-15] Various aspects of the equilibrium globular protein folding process, even in the framework of phenomenological theory, are still far from understood, and the most challenging problem in predicting the tertiary structure on the basis of the protein's primary structure remains unsolved.[15,16]

Computer simulations on model systems may be very helpful in elucidating the qualitative factors that specify the native structure. More exact methods, e.g., molecular dynamics applied to a realistic model of polypeptide chains, are feasible only in very limited cases (see, for example, Refs. 17–19). This is because of the large separation of the native and denatured states in phase space, the large number of intervening local minima on the free-energy surface, and the large number of degrees of freedom that must be treated. Thus, the simulation of the entire folding process by this method without *a priori* invoking some aspect of the native structure seems at present far beyond existing capabilities. Therefore, during the last decade, Monte Carlo (MC) methods have been applied to simplified models of real proteins in order to explore the equilibrium folding from the denatured to the native state and vice versa. The works of Go et al.,[20] Krigbaum and Lin,[21] Levitt,[22] Miyazawa and Jernigan,[23] and Segawa and Kawai[24] are representative. However, even in these idealized models it is necessary to introduce a specific biasing toward the native state in order to make the problem computationally tractable. Nevertheless, much useful information has been extracted from these simulations. For example, the relative role played by short- and long-range interactions in the dynamics of folding process, and the importance of some initial nucleation of secondary structure in the denatured state to the finding of the native structure, have been discussed.[17-21]

In this context it seems important to investigate models without employing a global potential and/or an *a priori* defined target native state, but rather where the model system is allowed to freely hunt over the entire configurational space to ultimately find the native state. It is hoped that this approach will give us some insight into the role of long- and short-range interactions as well as the influence of topological constraints on the folding equilibrium. Among the questions we address in present work are the following: What is the role of conformational stiffness (i.e., marginal secondary structure in denatured state) on the folding equilibrium? What makes the folding process more cooperative and what determines the all-or-none character of the transition? Since the direct approach to the problem is achieved at the expense of employing a simplified model, all the conclusions are qualitative. Nevertheless, we believe the model developed here possesses many features of the equilibrium folding process seen in real proteins and may be considered to be a useful zero-order approximation to it.

In the present work, we consider a very simple lattice model of a globular protein in which a hypothetical homoprotein chain is confined to the diamond lattice. Local conformational stiffness (which turns out to be responsible for

the formation of β-strands) is introduced by the slight *a priori* energetic preference for the *trans* (t) conformation over either of the two *gauche* conformations (g^+ or g^-). Furthermore, a nearest neighbor square-well potential is used in order to mimic the local nonspecific interactions within the model macromolecule. Additionally, we introduce a "long-range" coupling between the conformational states of nonbonded nearest neighbor chain segments separated by an appreciable distance down the chain contour; this latter effect is found to be required in order to reproduce some of the phenomena observed in the equilibrium folding process of globular proteins. In the present series of simulations, if the two nonbonded conformational states are t, these conformations are taken to be energetically stabilized. Physically, this represents the mutual stabilization of t conformational states by long-range interactions. The actual source of such long-range interactions may be due to hydrogen bonding, hydrophobic interactions between side chains, or salt-bridge formation. In the schematic model developed below, however, the long-range interactions are subsumed into a single energetic parameter ϵ_c; the purpose of the present series of simulations is to develop qualitative insight into the general nature of the folding process rather than to treat any specific protein. Finally, we might point out that, equivalently, one could bias the g^+ or g^- states, respectively, to produce right- or left-handed α-helices. Such dynamic MC simulations on the folding of α-helices will be undertaken in the near future. Here, we limit ourselves to the somewhat simpler case of modelling β-barrel proteins.

It should be pointed out that the present simulation model does not incorporate any specific biasing toward a particular "native" structure and therefore seems a well-defined equilibrium model in which the system itself choses the tertiary structure. The dynamics of the model system depends on the sampling method employed in MC process. As such it is rather artificial and thus not relevant to the dynamics of the folding process in real proteins.

METHOD

Description of the Model

The model macromolecule consists of n beads confined to a consecutive sequence of diamond lattice vertices. The mutual orientation of the $n - 1$ bonds requires specification of $n - 3$ conformational states. Three discrete conformations of every three-bond sequence are allowed, i.e., one t state and two gauche states. Thus, a single bond is represented as a vector of form $= [\pm 1, \pm 1, \pm 1]$ where the bond length is $|l| = 3^{1/2}$. We take ϵ_g as the energy difference between a t and either of the two g states. The hard-core part of local interactions between beads of the model chain is introduced by the *a priori* exclusion of multiple occupancy of lattice sites. The soft attractive part of the short-range site–site interaction is taken to be of the form of a square-well potential of depth ϵ_a, whose width extends only to the nonbonded nearest neighbors on the lattice [see Fig. 1(A and B)]. Additionally, the possibility of local coupling between the nearest neighbor nonbonded conformational states is described by the parameter ϵ_c and leads to the dependence of the attractive energy on the rotational degrees of freedom. The idea is

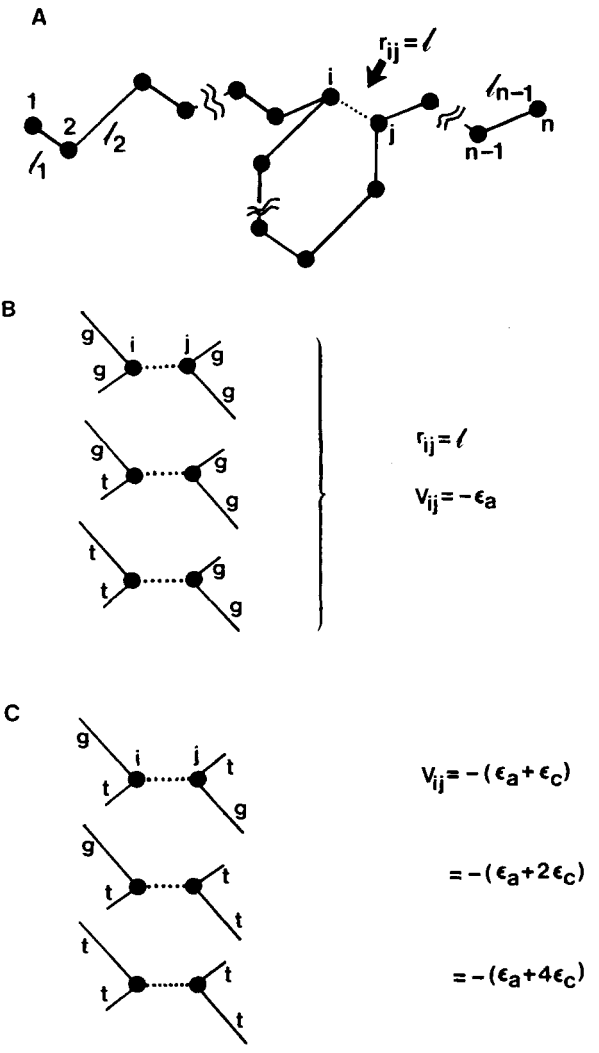


Fig. 1. Schematic representation of the conformation of a model macromolecule illustrating the coupling between the nearest neighbor interactions and the local chain conformation.

schematically depicted in Fig. 1(C) in which trivial permutations of $gt \rightleftarrows tg$ are omitted. Observe that the strength of the interaction V_{ij} is not *a priori* dependent on the mutual orientation of the two parts of the chain under consideration. Thus, the total configurational energy of a given conformation of the model chain is the following:

$$E = \sum_{i=2}^{n-2} (1 - f_{t,i}) \cdot \epsilon_g - \frac{1}{2} \sum_{i=1}^{n} \sum_{|j-i|>1}^{n} \delta(r_{ij} - l) \cdot \epsilon_a$$
$$- \frac{1}{2} \sum_{i=2}^{n-1} \sum_{|j-i|>1}^{n} \delta(r_{ij} - l) \cdot [f_{t,i-1} + f_{t,i}][f_{t,j-1} + f_{t,j}] \epsilon_c \quad (1)$$

where $f_{t,i} = 1$ for the t state of the rotational degree of freedom associated

with the ith bond ($f_{t,i} = 0$ for a g state or for end bonds, i.e., when $i = 1$ or $i = n - 1$), δ is the Dirac delta function, $r_{i,j}$ is the distance between the beads under consideration, $|\ |$ is the lattice spacing equal to the length of a bond, and $\frac{1}{2}$ is the symmetry factor. According to the above definition of the configurational energy, the reference state (with energy $E_{ref} = 0$) corresponds to an isolated chain in the fully extended all-*trans* conformation. Another possible choice of reference state is the self-avoiding random coil at infinite temperature. In the last case $E_{ref} \neq 0$, provided that at least one of the energetic parameters—ϵ_g, ϵ_a, or ϵ_c—is nonzero.

While in principle ϵ_g, ϵ_a, and ϵ_c could be site dependent, in the simulations described below we assume that they are site independent for several reasons. First, in exploring the properties of this class of models we wish to keep the number of adjustable parameters to a minimum. Second, and of greater importance, is our wish to explore whether or not relatively unique tertiary structures can be formed in the absence of site-specific interactions, and if so, what are the requirements for the formation of such structures? Thus, we view these simulations as computer experiments designed to help elucidate the general rules of protein folding.

Sampling Procedure

In order to estimate the average properties of the model macromolecule at a given temperature, a dynamic MC method was used to generate a long sequence of equilibrium states.[25] The following set of elementary micromodifications of the chain conformation have been employed:

1. Reptationlike motion of the chain backbone involving the clipping-off of one segment at one of the chain ends, which is then added to the opposite end in a randomly selected direction [see Fig. 2(A)].
2. Random reorientation of the two bonds at the chain ends [see Fig. 2(B)].
3. Three-bond kink motion [see Fig. 2(C)].
4. Four-bond kink motions that create new conformations in the inner part of the chain [Fig. 2(D)].

The sequence of attempts to make a particular kind of motion was selected by a random method. For some computations, the entire set of elementary moves listed above have been employed. However, because reptationlike motions are the most efficient at sampling phase space,[26,27] the majority of the results discussed below have been obtained by simplified algorithms involving only reptation moves and chain end reorientation. The direction of the slithering of the chain down its contour was kept unchanged until the first unsuccessful move, after which the slithering was switched to the opposite direction. The motion of chain ends was randomly mixed with reptation.

Every micromodification was accepted according to the standard Metropolis criterion. If E_{old} is the configurational energy of the system in a given state and E_{new} is the energy of the new trial one, then the probability of acceptance of an attempted move is given by

$$p_{new/old} = \min\{\exp[-(E_{new} - E_{old})/k_B T], 1\} \tag{2}$$

Thus, the distribution of energies of the generated states tends to a Boltz-

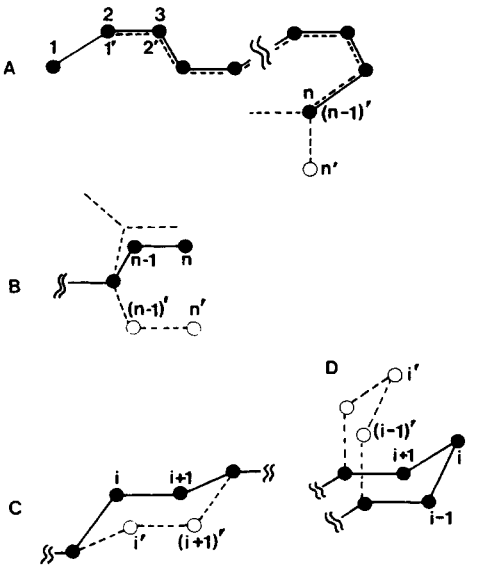


Fig. 2. The elementary modifications of the chain conformation on the diamond lattice employed in the MC algorithm.

mann distribution, and therefore the ensemble average of a quantity could be estimated from the simple arithmetic mean. For example, the mean energy of the system $\langle E \rangle$ is computed as

$$\langle E \rangle = M^{-1} \sum_{i=1}^{M} E_i \tag{3}$$

where M is the number of cycles of MC algorithm. Similarly, the average fraction of t states, $\langle f_t \rangle$, follows from

$$f_t = M^{-1}(n-3)^{-1} \sum_{i=1}^{M} \sum_{j=2}^{n-2} f_{t,j} = (n-3)^{-1} \sum_{j=2}^{n-2} \langle f_{t,j} \rangle \tag{4}$$

and the average number of nearest neighbor nonbonded pairs (which is equivalent to the frequency of occupation of the sites in the first coordination sphere by other nonbonded beads for the lattice under consideration), $\langle \nu \rangle$, is

$$\langle \nu \rangle = 2M^{-1}[(q-2)n+2]^{-1} \sum_{m=1}^{M} \sum_{i=1}^{n} \nu_{m,i} \tag{5}$$

where $\nu_{m,i}$ is the number of bead—bead interactions at site i in the mth chain. $(q-2)n+2$ is the coordination number of the macromolecule, with $q=4$ the coordination number of the diamond lattice. Very long sequences of states have been generated (i.e., M is large) for the purpose of estimating the various average properties. For the smallest systems, the number of MC iterations was in the range of 10^6 and for the largest systems M ranged from

10^7–10^8, with the particular value of M depending on the algorithm and temperature.

The question arises as to whether the above procedure, employing only reptationlike steps, is ergodic, and whether the algorithm really samples the entire configurational space. The following facts suggest that the method employed here leads to a "correct" estimation of the equilibrium properties. First, there seems to be no difference between the results (based upon tests on smaller systems) obtained with various sets of elementary motions. Second, the cooling–heating sequences of computational experiments give statistically the same results, independent of the initial state and the seed used in the random number generator. Furthermore, the flow charts for very long runs do not exhibit any systematic trends, except for the oscillations (discussed later), which are typical of two-state models in the vicinity of the transition temperature. However, one must exercise some caution concerning the results of the sampling efficacy at temperatures well below the observed transition to the globular state. Subsequent to the formation of the native structure, the entire set of elementary motions listed above becomes less and less effective in sampling phase space. Hence, at low temperatures we cannot exclude the possibility of more substantial changes occurring in the configuration of the system than those observed in the present computer experiments. In other words, the algorithm may exhibit some deviation from ergodicity well below the transition temperature due to the slowing down of the evolution of the system, which is presumably related to the increasing local free-energy barrier in configurational space. It also should be noted that a reptationlike procedure can be safely used only in the absence of differentiation of the particular segments of the chain molecule—i.e., every bead must be identical. This condition holds in the present model of a homopolymeric molecule. The study of the folding process of nonhomopolymeric protein models with a specific primary structure requires an algorithm that allows for an efficient change of register of interactions between the various kinds of protein segments. The details of this calculation will be reported in the near future.

RESULTS

The Collapse of Semiflexible Polymers with $\epsilon_c = 0$

The results of a study of the collapse transition of semiflexible polymers ($\epsilon_g \neq 0$ and $\epsilon_a \neq 0$) without cooperative interactions (model A) have been recently reported.[28,29] Although the observed collapse to a well-defined ordered structure resembling a β-barrel[16] has many features in common with the equilibrium folding process of globular proteins, there are also some major differences. First of all, the denatured state of the model system exhibits too much ordering on the local scale just prior to the transition to the native structure in comparison with that seen in real protein systems.[15,30] Moreover, only in the case of significant stiffness prior to the transition does the low-temperature state exhibit specific global ordering in the form of a β-barrel. These polymers exhibit an increase of the random-coil dimensions with decreasing temperature, and in this case the sharp collapse transition seems

consistent with the two-state model.[2-6] However, if the flexibility of the model chain exceeds a critical value, the collapse is smooth and second order, with a crossover regime whose width depends on the degree of polymerization and is poorly represented by a two-state model.

From these previous studies, it appears evident that an important aspect of protein folding, presumably related to the cooperative character of intramolecular interactions, must be added to the model. Therefore, the present simulation focuses on the effect of coupling between conformational stiffness (or secondary structure) and the strength of the long-range interactions between the chain units. In this way, it is hoped that the more cooperative character of the folding process can be taken into account.

The System with Only Cooperative Interactions: $\epsilon_g = \epsilon_a = 0$, $\epsilon_c \neq 0$

Model B has no intrinsic local conformational stiffness; i.e., in the absence of long-range interactions, all the rotational isomers (g^+, g^-, and t) are *a priori* equally weighted; however, the excluded-volume effect (hard core site–site repulsions), with the resultant expansion of the chain in the athermal system, produces a slight biasing towards the trans conformation. The only "soft" attractive interactions result from the coupling between isomeric states of adjacent nonbonded pieces of the chain [see Fig. 1(C)]. Hence, on setting $\epsilon_a = 0$, $\epsilon_g = 0$, and $\epsilon_c \neq 0$, the possibility of long-range intrachain interactions arises only from cooperative interactions in the Hamiltonian [see Eq. (1)] of the system. In this sense the present case is opposed to that for semiflexible polymers (model A) with both ϵ_g and ϵ_a nonzero, but with $\epsilon_c = 0$. Therefore, one might expect to see qualitatively different features in the collapse transition.

Figure 3 shows the behavior of various properties of the model as a function of the reduced temperature $k_B T/\epsilon_c$. In Fig. 3(A), the expansion factor α_s^2 of the mean-square radius of gyration $\langle S^2 \rangle$ with

$$\alpha_s^2 = \langle S^2 \rangle / \langle S_0^2 \rangle \tag{6}$$

for chains of length $n = 200$ (400) in the open (solid) diamonds is presented. $\langle S^2 \rangle$ is defined by

$$\langle S^2 \rangle = \frac{1}{2n^2} \sum_{i,j=1}^{n} \sum r_{ij}^2 \tag{7}$$

and $\langle S_0^2 \rangle$ is the ideal rotational isomeric state value. Using the Flory method,[31] $\langle S_0^2 \rangle = 197.8$ for the $n = 200$ case and $\langle S_0^2 \rangle = 397.8$ for the $n = 400$ case, respectively. One can observe a sharp transition from a voluminous random coil to a dense globular state.

As shown in Fig. 3(B), the average fraction of trans states in the chain is almost temperature independent over a wide range of $k_B T/\epsilon_c$ and then increases sharply when the collapse occurs. Observe, however, that there is still a substantial fraction of g bonds in the dense globular state. Furthermore, as shown in Fig. 3(C), the average number of nearest neighbor pairs,

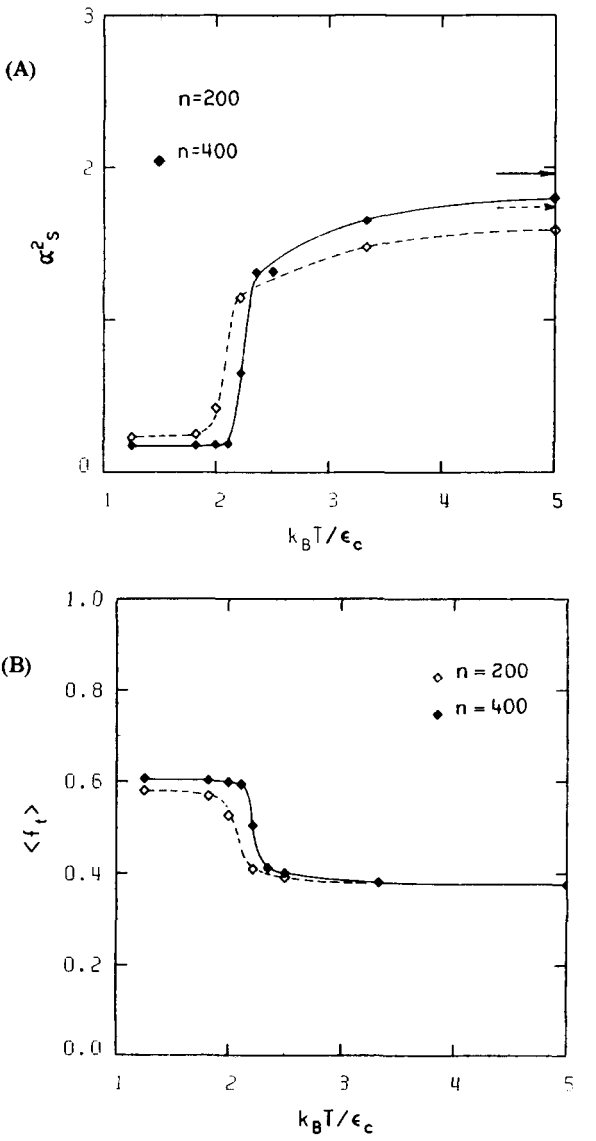


Fig. 3. Temperature (k_BT/ϵ_c) dependence of the average chain properties for model B, $n = 200$ (open symbols) and $n = 400$ (closed symbols). (A) Expansion factor $\alpha_s^2 = \langle S^2 \rangle / \langle S_0^2 \rangle$; the arrows indicate the results for a self-avoiding random walks (SAWs; see ref. 37). (B) Mean fraction of trans states $\langle f_t \rangle$. (C) Mean number of nearest neighbor pairs per bead, $\langle \nu \rangle$.

$\langle \nu \rangle$, (contacts of nonbonded beads) also increases rapidly in the vicinity of the transition. The transition, although very sharp, seems continuous for any reasonable number of polymer segments n; the transition temperature increases with increasing n in the range of chain length under consideration. We have not analyzed the asymptotic $n \to \infty$ character of the transition.

The first conclusion that can be drawn from the comparison of the two models (model B with $\epsilon_a = \epsilon_g = 0$, and $\epsilon_c \neq 0$ vs model A with $\epsilon_c = 0$, $\epsilon_a \neq 0$, and $\epsilon_g \neq 0$) is that the inclusion of cooperative interactions alone does not

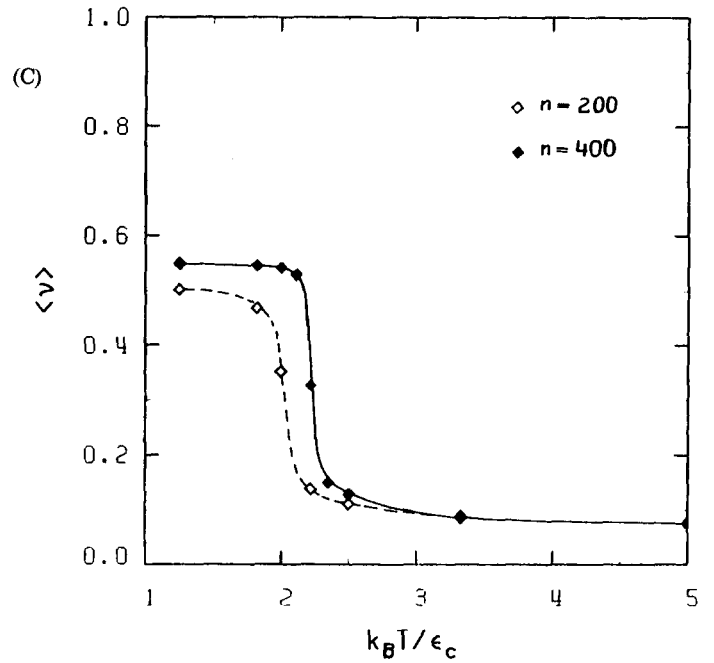


Fig. 3. *(Continued from the previous page.)*

lead to any secondary structure (e.g., an appreciable fraction of isolated stretches) in the high-temperature, random-coil state. In the denatured state in model B, unlike in model A, there is no increase of $\langle f_t \rangle$ during the "cooling" of the system over a wide temperature range. Furthermore, in model B the coil dimensions $\langle S^2 \rangle$ decrease rather than initially increase with decreasing temperature as in model A. The transition is considerably steeper than that for the flexible polymers ($\epsilon_g = 0$, $\epsilon_a = 0$)[32-34]; however, it is not so sharp as that for stiff polymers ($\epsilon_a \neq 0$, $\epsilon_g \neq 0$) with $|\epsilon_g/\epsilon_a| > 2$ and a moderate number of segments. Since there is no significant reduction of the available conformational space in the random-coil state of model B, the degree of the order induced by the collapse is not sufficient to produce a unique tertiary structure in the globular state. Finally, as seen in the snapshot projections of representative configurations of chains with $n = 100$, $n = 200$, and $n = 400$, respectively, in Fig. 4(A–C), the extent of ordering in the collapsed state is local and rather moderate. That is, there is no well-defined unique native state. In Table I, we compare some properties of the folding equilibrium observed in model A and model B, respectively.

The above results show that cooperative long-range interactions in the absence of intrinsic local chain stiffness can produce a sharp transition from an essentially random-coil state to a dense globular state having marginal local order, whereas globular proteins have an essentially unique tertiary structure. Therefore, it appears that a model incorporating both local configurational stiffness as well as cooperative interactions might exhibit more features of the conformational transition of globular β-proteins. Based on models A and B, the first property, local stiffness, operating in conjunction

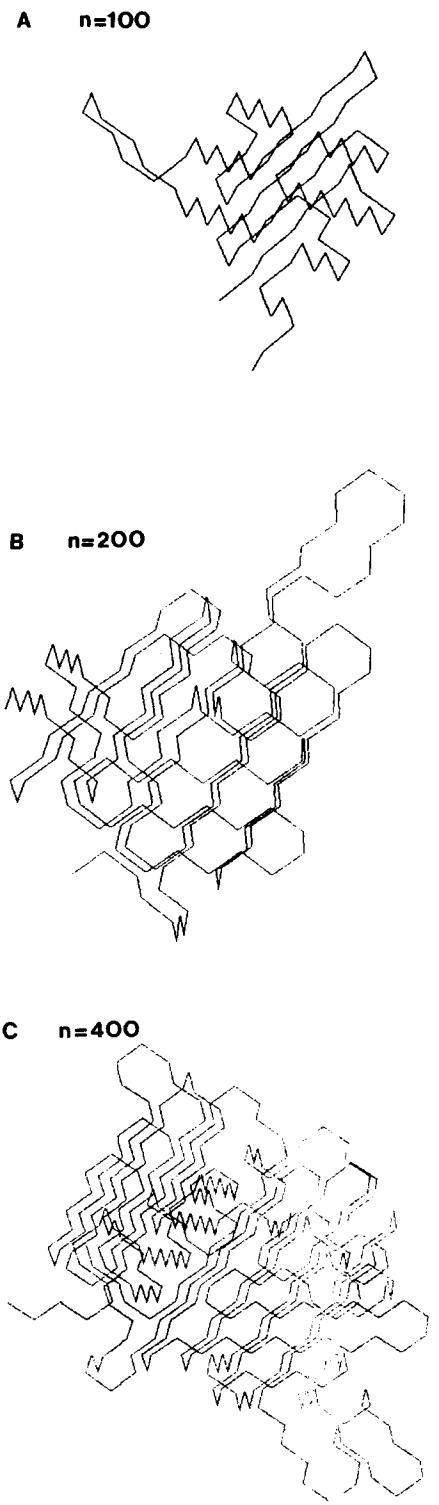


Fig. 4. Snapshot configurations of the collapsed structure for Model B ($\epsilon_a = 0$, $\epsilon_g = 0$, $\epsilon_c \neq 0$) for various chain lengths. (A) $n = 100$, (B) $n = 200$, (C) $n = 400$.

TABLE I
Comparison of Properties of the Folding Process in Models A and B

Properties at transition	Model A $\lvert\epsilon_a/\epsilon_g\rvert = 1/4$, $\epsilon_c = 0$		Model B $\epsilon_a = \epsilon_g = 0$, $\epsilon_c \neq 0$	
	$n = 200$	$n = 400$	$n = 200$	$n = 400$
ϵ_a/k_BT or ϵ_c/k_BT	0.55	0.50	0.5	0.45
$\langle f_t \rangle^a$	0.818	0.786	0.409	0.401
$\langle f_t \rangle^b$	0.911	0.900	0.568	0.594
$\langle \nu \rangle^a$	0.0373	0.0395	0.1372	0.1485
$\langle \nu \rangle^b$	0.4643	0.4991	0.4676	0.5285
$\langle n_t \rangle^{a,c}$	10.4	8.5	2.41	2.34
$\langle n_t \rangle^{b,c}$	35.1	34.6	4.90	5.52

[a] Random-coil state, just prior to the transition (or at the transition).
[b] Collapsed state, just after (or at the) transition.
[c] Weighted average length of t stretches.

with intrachain cooperativity, seems responsible for the substantial ordering of the globular state. This kind of model (model C) is analyzed in the next section.

Semiflexible Chain Model with Cooperative Interactions

The Character of the Transition

In the following we focus on the properties of our candidate model of a globular protein—model C—which should possess all the qualitative features of protein folding; namely, the case where both ϵ_c and ϵ_g are nonzero. Before discussing the effect of varying the relevant parameters (chain length, stiffness, cooperative interactions, and noncooperative attractive interactions) on the properties of the random-coil denatured state, and the globular native structure, we begin by analyzing, in detail, the character of the transition (or folding equilibrium) itself. This is done for the case of a chain having $n = 200$ with a degree of intrinsic stiffness $\epsilon_c = \epsilon_g/4$ and without any attractive interactions between nonbonded nearest neighbors ($\epsilon_a = 0$). In Fig. 5(A–E) the values of $\langle S^2 \rangle, \langle f_t \rangle, \langle \nu \rangle$, the heat capacity C_v, and the energy E are plotted against the reduced temperature of the system defined as k_BT/ϵ_g. The solid diamonds correspond to the average results obtained from a series of computer experiments involving cooling–heating–cooling sequences, and the solid lines are drawn by an arbitrary interpolation through the solid diamonds. For the case of $\langle S^2 \rangle$ in Fig. 5(A), we also plot the "experimental" results obtained from a single cooling (or/and heating) sequence in the open diamonds. For every property under consideration, a narrow transition region is observed. In the vicinity of the apparent transition temperature T_t, the data exhibit large scattering, with the characteristic features of a two-state model. Namely, we find for rather short runs (e.g., the data shown in the open diamonds for $\langle S^2 \rangle$) that the observed measurements exhibit a bimodal distribution. Most of the observed properties are typical of that seen in the low-temperature globular state or in the high-temperature random state. Of course, in the limit of a very long MC run, one should observe a smooth

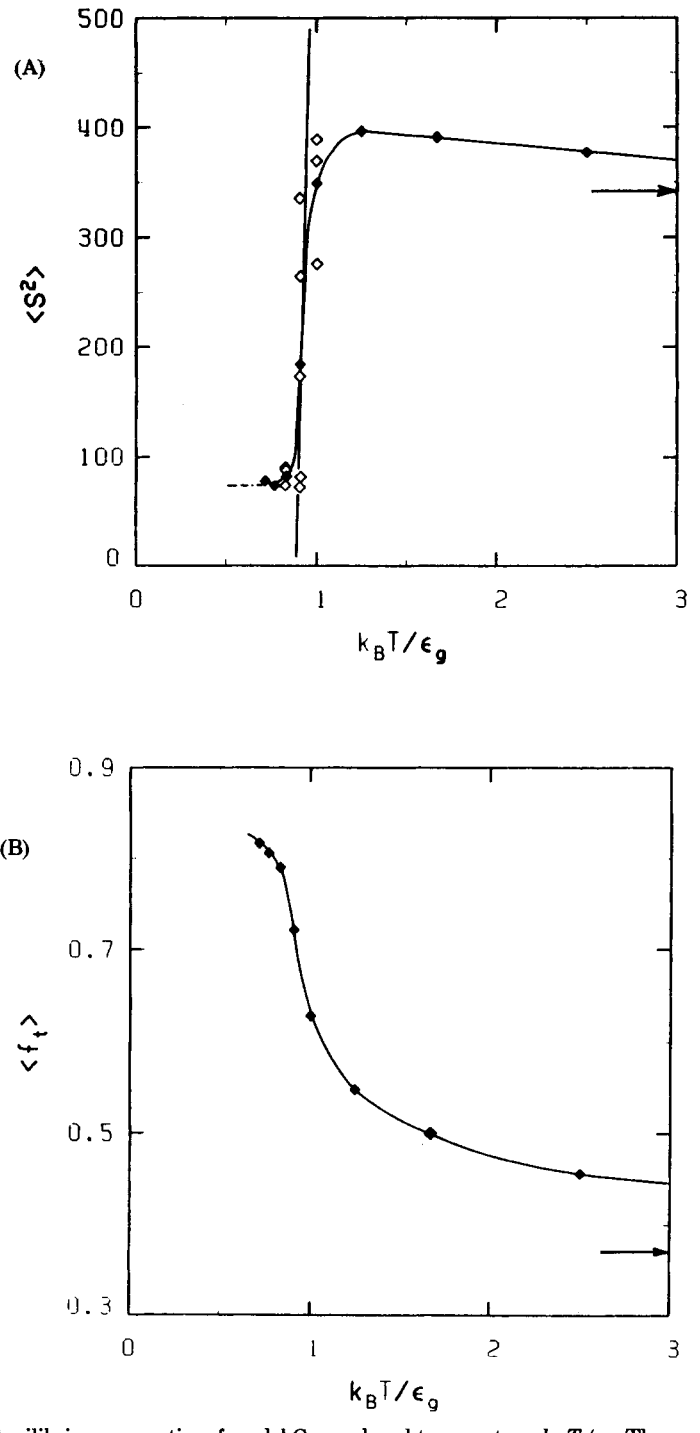
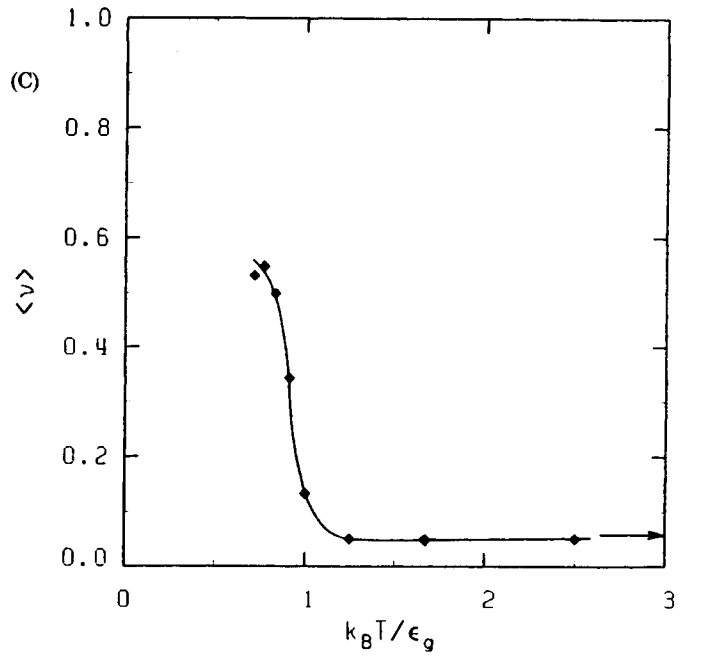
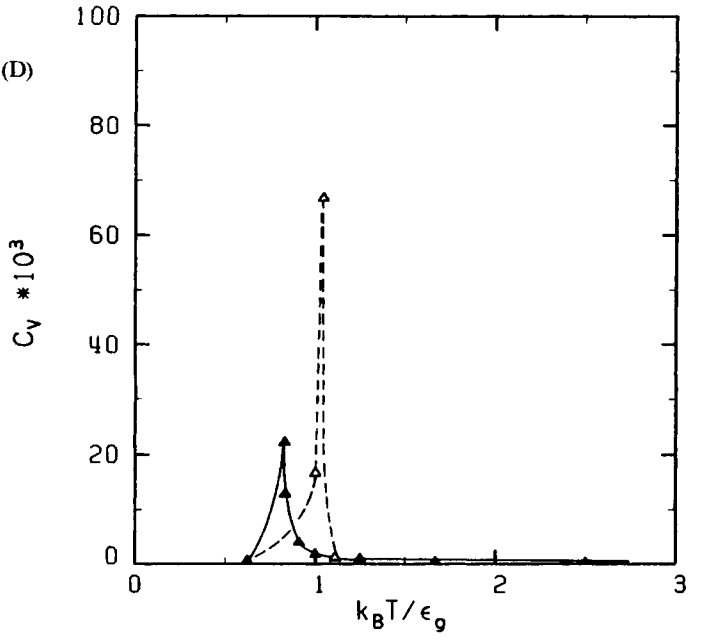


Fig. 5. Equilibrium properties of model C vs reduced temperature k_BT/ϵ_g. The case of $\epsilon_a = 0$, $\epsilon_c = \epsilon_g/4$ and $n = 200$. (A) The mean-square radius of gyration $\langle S^2 \rangle$. The open diamonds show the results obtained from shorter runs in the crossover regime. The solid diamonds correspond to averages over several runs (in the high-temperature range, the scattering is below size of the symbols). The arrow shows the value of $\langle S^2 \rangle$ for $T \to \infty$ (i.e., the SAW case). Note the change of definition of the reduced temperature in comparison to model B. (B) Average value of $\langle f_t \rangle$ vs reduced temperature k_BT/ϵ_g. (C) The average values of $\langle \nu \rangle$ vs temperature (the arrow indicates the SAW value in the limit that $T \to \infty$). (D) Heat capacity of the system per bead, $C_V = (\langle E^2 \rangle$





$-\langle E\rangle^2)/(nk_B^2T^2)$. The solid triangles correspond to a single run in a sequence of decreasing temperature, and the open triangles are obtained from a heating sequence. The low-temperature peak is at $k_BT/\epsilon_g = 0.83$ and the high-temperature peak is at $k_BT/\epsilon_g = 1.04$. Both should coincide in the limit of an infinitely long MC run. The estimated reduced transition temperature is $k_BT_t/\epsilon_g = 0.94$. (E) Plot of reduced energy per bead $E/\epsilon_g n$ vs reduced temperature k_BT/ϵ_g, for $n = 200$, $\epsilon_a \equiv 0$, and $\epsilon_c = \epsilon_g/4$. The solid diamonds correspond to ensemble averages, and the open diamonds are the results from shorter subruns where no transitions from the native to the denatured state or vice versa are observed. The estimated calorimetric energy change of transition is $\Delta E/\epsilon_g n = 0.64$.

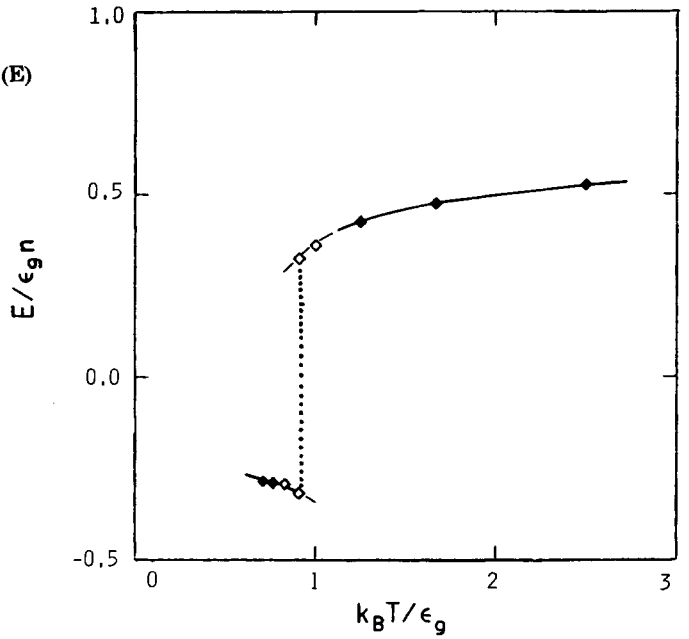


Fig. 5. *(Continued from the previous page.)*

crossover with the change of temperature; however, essentially two states contribute to the observed averages (whose relative weights are temperature and model dependent).

The plot of energy of the system E vs reduced temperature given in Fig. 5(E) can be used to estimate the energy difference between the native and denatured states, ΔE, which is equivalent to the "calorimetric" energy of denaturation. Assuming that the values of E obtained from shorter runs (when no transition from a low-temperature state to the high-temperature state or vice versa takes place) may be considered as a reliable extrapolation of the energy of both states, one obtains in the vicinity of the transition that $\Delta E/\epsilon_g n = \Delta E_{cal}/\epsilon_g n \cong 0.64$.

Let us consider the equilibrium between hypothetical native and denatured states

$$\text{native } (N) \rightleftarrows \text{denatured } (D)$$
$$K = f_D/(1 - f_D) \tag{8}$$

where f_D is the fraction of molecules in the denatured state at equilibrium. From the Van't Hoff equation, one obtains

$$d(\ln K)/dT = \frac{1}{f_D(1 - f_D)} \frac{df_D}{dT} = \frac{\Delta E_{VH}}{k_B T^2}$$

where ΔE_{VH} is the standard energy change (per molecule) going from the native (N) to the denatured state (D). In dimensionless units, the following

relationship is easily derived:

$$df_D/d(k_BT/\epsilon_g) = f_D(1 - f_D) \cdot \frac{\Delta E_{VH}}{\epsilon_g}\left(\frac{\epsilon_g}{k_BT}\right)^2$$

If the transition has an all-or-none character, then the calorimetric energy should be exactly the same as measured from temperature dependence of equilibrium constant, i.e., $\Delta E_{cal} = \Delta E_{VH}$,[35,36] and essentially any property of the system could be used to estimate the equilibrium constant. Let us consider the temperature dependence of radius of gyration $\langle S^2 \rangle$ [Fig. 5(A)] as an example. At $k_BT/\epsilon_g = 0.909$, which is in the transition crossover regime, $\langle S^2 \rangle = 183.6$. Assuming that two states contribute to the measured average, one obtains $f_D \cong 1/3$. Thus the slope $df_D/d(k_BT/\epsilon_g)$ should be approxi-

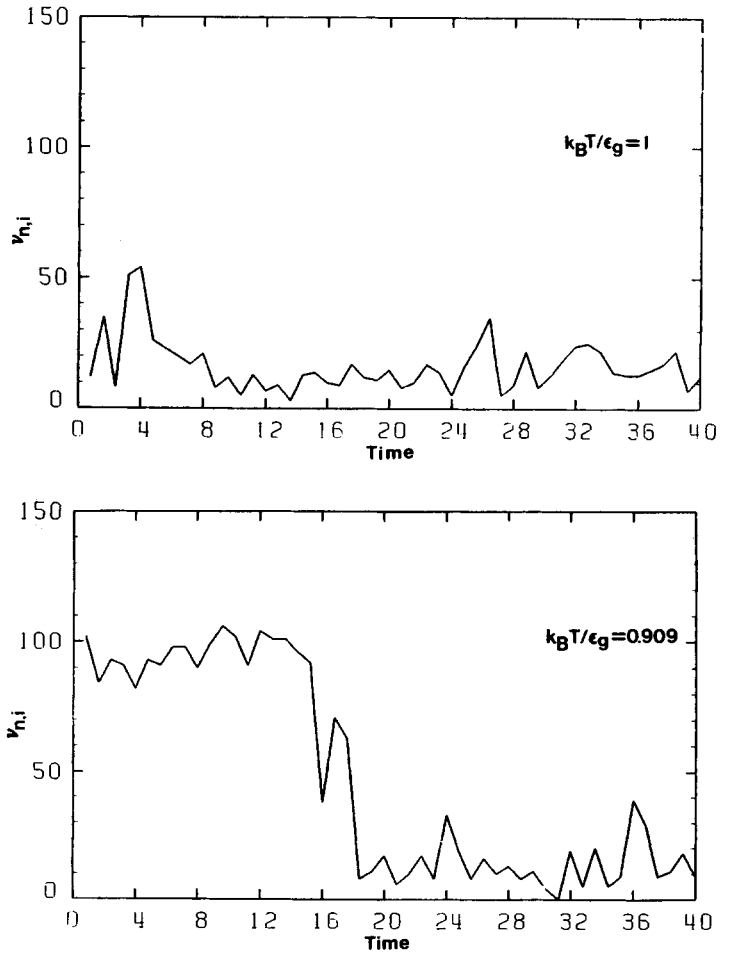


Fig. 6. Single-run flow charts for the number of nearest neighbor pairs of segments in the chain, $\nu_{n,i}$, in the vicinity of the transition temperature. (A) An example of a run above the transition temperature, $k_BT/\epsilon_g = 1.0$; (B) at the transition temperature, $k_BT/\epsilon_g = 0.909$; and (C) below the transition temperature, $k_BT/\epsilon_g = 0.833$. The numbers on the abscissa correspond to successive "photographs" of the molecule, taken every $2.5 \cdot 10^5$ iterations.

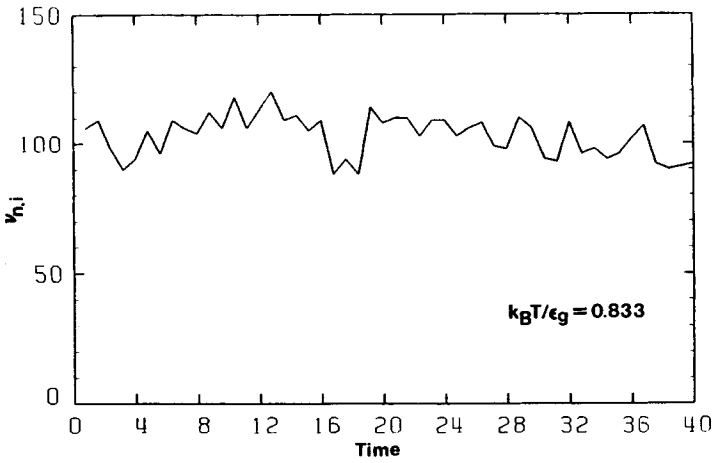


Fig. 6. *(Continued from the previous page.)*

mately equal to 34.4. Since $d\langle S^2\rangle = (\langle S^2\rangle_{\text{denatured}} - \langle S^2\rangle_{\text{native}}) \cdot df_D = 263 \cdot df_D$, the slope of $\langle S^2\rangle$ vs k_BT/ϵ_g should be $9.05 \cdot 10^3$; this is the value of the slope of the line drawn in Fig. 5(A) in the range of $\langle S^2\rangle$ between 100 and 300. Within the error of the sampling, the transition is consistent with the two-state model of protein folding.[2-6] This behavior is also exhibited by the other properties studied in the transition region.

Additional and perhaps more compelling evidence for the two-state equilibrium character of the observed folding process comes from the trajectory analysis of the system's properties; flow charts of various properties of the model can be used for this purpose. In Fig. 6(A–C), the flow charts of the number of nearest nonbonded neighbors (per chain) are compared for three temperatures that are close to each other. At the highest temperature, $k_BT/\epsilon_g = 1.0$ [Fig. 6(A)], one can observe a small fluctuation of $\nu_{n,i}$ around a low average value, which is the proper one for a quasirandom-coil state. At the lowest temperature, $k_BT/\epsilon_g = 0.83$ [Fig. 6(C)], there are again relatively small fluctuations around the value typical of a high-density globular state. At the intermediate temperature, $k_BT/\epsilon_g = 0.91$ [Fig. 6(B)], which is very close to the transition temperature, one can observe a period of low values or high values, within which are small oscillations about the mean that are then punctuated by sudden jumps between the two mean values; this corresponds to the dynamic equilibrium about the mean between the two states—denatured and native. For sufficiently long molecules (in the range of $n = 100$ and larger, depending on the degree of cooperativity assumed) the transition has an essentially all-or-none character. For shorter chains, the two-state behavior gradually washes out since the degree of uniqueness of the collapsed state becomes less. Interestingly, there is also a decreasing uniqueness of the high-density structure in the opposite limit of very long polymers; however, in that case the transition remains very sharp in spite of the possibility of formation of multiple domains in the low-temperature state.

Taking the above observations into account, the transition temperature can be estimated as that point where the slope of the mean values of the various properties vs reduced temperature assumes its largest value, and this should

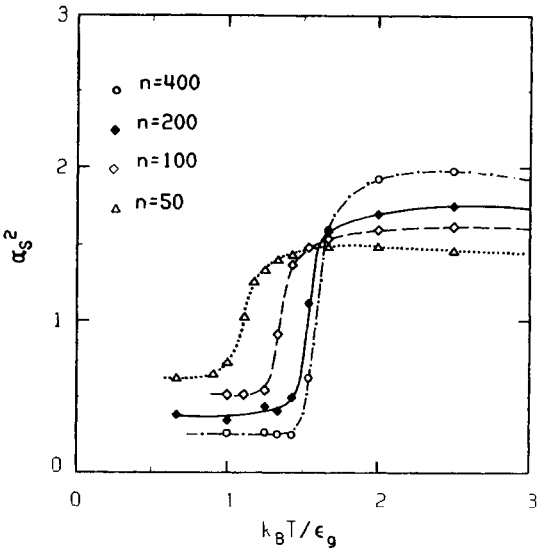


Fig. 7. The temperature dependence of the expansion factor α_s^2, for model C, with $\epsilon_a = 0$ and $\epsilon_c = \epsilon_g/2$ at various chain lengths n.

be very close to the temperature where the contributions to the average of the native and denatured states are equal. Direct application of the last criterion, however appealing, was not used due to the very large computational cost, when n is large, of preparing an equilibrium ensemble of mutually independent states close to the transition. Another complication could emerge from the artificial dynamics of the model system. For the latter reason, we always performed several cooling–heating–cooling experiments in the vicinity of the collapse.

Effect of Chain Length and Degree of Cooperativity on the Folding Equilibrium

A measure of the degree of cooperativity introduced into the model is the value of the ratio ϵ_c/ϵ_g, where ϵ_g/k_BT is taken to be the independent variable. At a fixed ϵ_c/ϵ_g, we examined the effect of the chain length on the folding equilibrium. Representative behavior is displayed in Fig. 7, where the expansion factor α_s^2 is plotted vs reduced temperature, k_BT/ϵ_g, for various values of n at fixed $\epsilon_c/\epsilon_g = 1/2$. The reference state of α_s^2 is an ideal chain at infinite temperature (or $\epsilon_g = 0$), i.e., a tetrahedral lattice RIS model.[31] We found that the transition becomes steeper with increasing n. However, the uniqueness of the folded state is greatest in a loosely defined range of moderate values of n. The transition temperature migrates slowly towards higher values with increasing chain length; presumably there is a limiting value as $n \to \infty$.

Similar behavior was observed when the cooperativity is lowered (i.e., for the chain of higher intrinsic stiffness). This is shown in Fig. 8, which differs from the previous figure in that the value of ratio $\epsilon_c/\epsilon_g = 1/4$ rather than $1/2$. Further increase of the chain stiffness leads to behavior very similar to

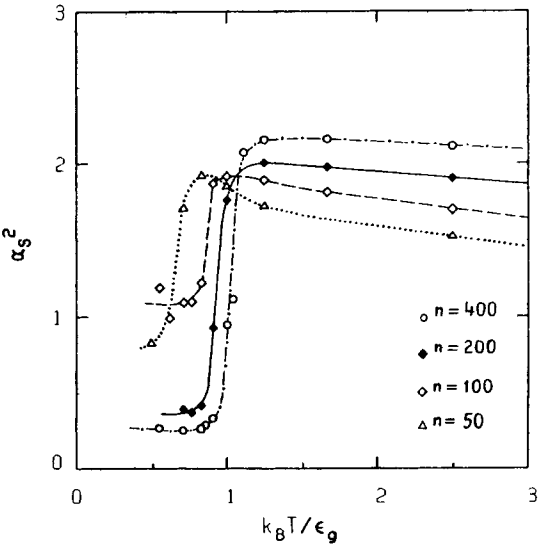


Fig. 8. The temperature dependence of the expansion factor α_s^2 for model C, with $\epsilon_a = 0$ and $\epsilon_c = \epsilon_g/4$ at various chain lengths n.

that exhibited by semiflexible polymers (model A, see section on the collapse of semiflexible polymers). As mentioned above, that model, with $\epsilon_c = 0$, $\epsilon_a \neq 0$, and $\epsilon_g \neq 0$, possesses too much secondary structure (β-strands) in the high-temperature phase, just prior to the transition. As demonstrated in Fig. 9, where the fraction of t states in various models is plotted against reduced

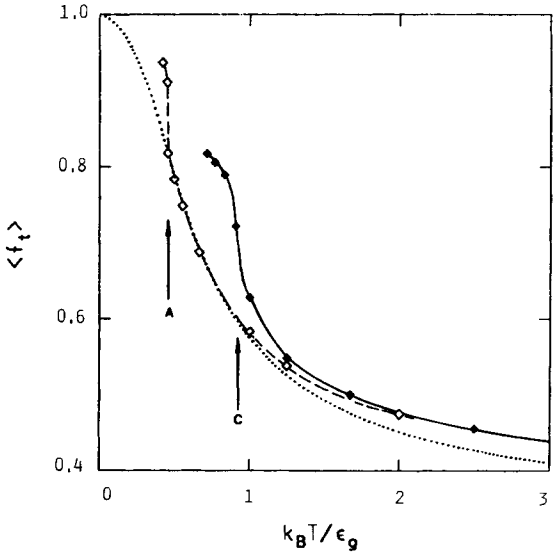


Fig. 9. Comparison of the degree of stiffness reached prior to the transition in model A ($\epsilon_a = \epsilon_g/4$), open diamonds, and in model C ($\epsilon_a \equiv 0$, $\epsilon_c = \epsilon_g/4$), solid diamonds, respectively. $\langle f_t \rangle$ is plotted for the case of $n = 200$. The arrows identified by A and C denote the reduced transition temperature in models A and C, respectively. The dotted line represents the ideal RIS model results.

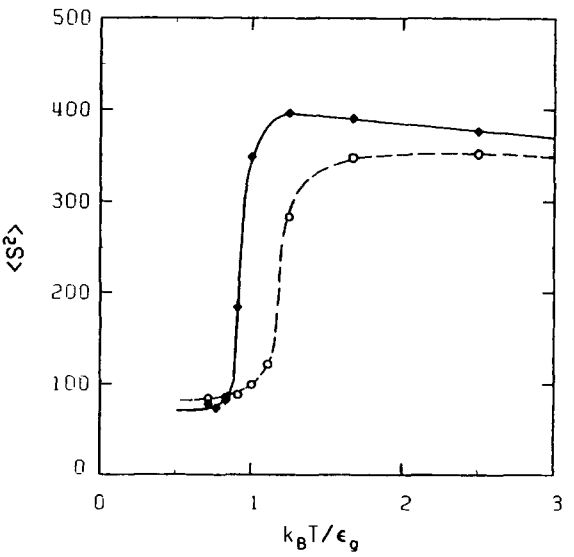


Fig. 10. Effect of an attractive interaction on the collapse in model C with $\epsilon_a = 0$ and $\epsilon_a = \epsilon_c$, respectively, in the solid diamonds and open circles. In both cases, $\epsilon_c = \epsilon_g/4$ and $n = 200$.

temperature, this disadvantage has been eliminated in model C ($\epsilon_c \neq 0$). The dotted curve represents the prediction for a RIS model without excluded volume, the open diamonds show f_t from model A with $\epsilon_a/\epsilon_g = 1/4$ and $\epsilon_c = 0$, and the solid diamonds depict f_t from model C with $\epsilon_a = 0$, $\epsilon_c/\epsilon_g = 1/4$. One can note that the chain stiffness reached just prior to the transition in the case of $\epsilon_c \neq 0$ ($f_t = 0.628$) is substantially smaller than that of previous model with $\epsilon_c = 0$ ($f_t = 0.818$). Thus model C possesses features similar to real β-proteins, where the degree of secondary structure of the denatured state is usually rather small. Therefore, we conclude that cooperative interactions are essential to reproduce the globular protein chain folding phenomena and can be qualitatively modeled in a very simple way. Some initial ordering, however, seems required to produce a well-defined tertiary structure. This can be achieved by introducing a small degree of intrinsic stiffness into the model. These locally ordered stretches appear to form the nuclei around which the well-defined globular structure forms, i.e., the system must have a predilection to form a β-sheet in order to form the unique globular structure. The presence of a net noncooperative attractive interaction (ϵ_a) plays a rather marginal role in model C. However, a relatively strong attractive interaction can change some aspects of the details of the folding equilibrium quite a bit. An example is shown in Fig. 10, where $\langle S^2 \rangle$ is plotted against reduced temperature $k_B T/\epsilon_g$, for the case of $\epsilon_a \equiv 0$ (solid diamonds) and $\epsilon_a = \epsilon_c = \epsilon_g/4$ (open circles), respectively, with the same $n = 200$. Not surprisingly, the inclusion of an attractive interaction shifts the folding equilibrium toward higher temperatures. As a result, the fraction of t states, and the coil dimensions, in the high-temperature state are somewhat smaller just prior the transition. However the properties of the low-temperature, native state remain almost completely unaffected.

TABLE II
Transition Temperatures k_BT/ϵ_g for Model C for Various Chain Lengths and Conditions

Model Parameters	$n = 50$	$n = 100$	$n = 200$	$n = 400$
$\epsilon_a \equiv 0$, $\epsilon_c = \epsilon_g/2$	1.11	1.33	1.53	1.60
$\epsilon_a \equiv 0$, $\epsilon_c = \epsilon_g/4$	0.65	0.87	0.94	1.05
$\epsilon_a \equiv 0$, $\epsilon_c = \epsilon_g/8$	—	—	0.67	—
$\epsilon_a = \epsilon_c = \epsilon_g/4$	—	—	1.18	—

The transition temperatures for various systems under consideration (model C) are summarized in Table II.

The Structure of the Low Temperature State

When the cooperative conformational intrachain interactions are included with some intrinsic stiffness (which models the "secondary" structure) we obtain a well-defined native structure after the collapse transition. There are, however, two caveats. First, semiflexible polymers (with cooperative interactions) without any site-specific flexibility (specific secondary structure) down the chain tend to give a constant length of β-strands (which depends on the ratio ϵ_g/ϵ_c). Therefore, very short polymers with a very low (2 or 3) number of β-strands in the ordered structure have too small an interaction energy to stabilize this structure. This is one of the reasons why, in the short polymer

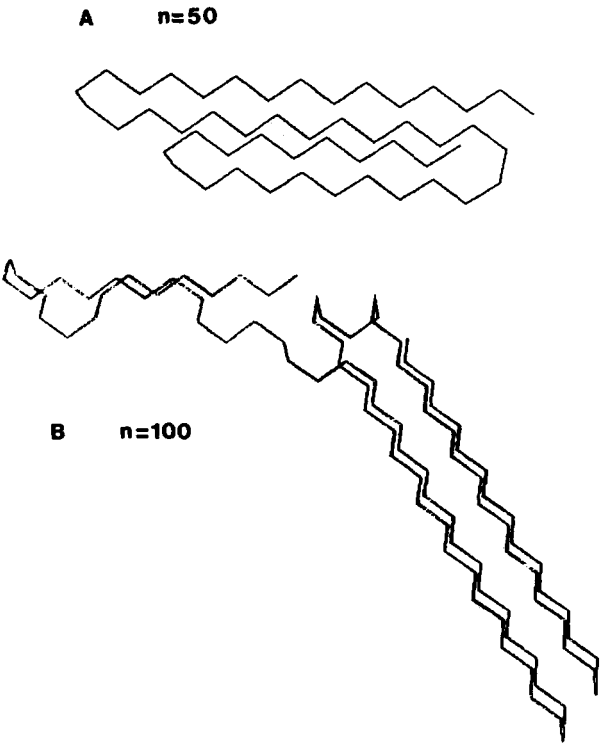


Fig. 11. Representative configuration of the low-temperature states of model C ($\epsilon_a \equiv 0$, $\epsilon_c = \epsilon_g/2$). (A) $n = 50$, (B) $n = 100$, (C) $n = 200$, (D) $n = 400$.

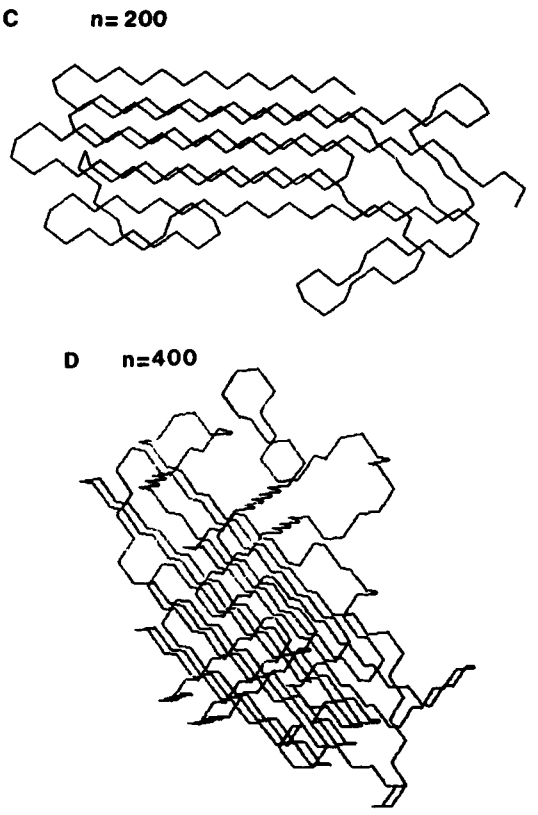


Fig. 11. *(Continued from the previous page.)*

limit, the two-state picture of equilibrium folding washes out. The second caveat is relevant to the case of long polymers (say $n \geq 400$) where a large number of β-strands leads to the possibility of the arrangement of the low-temperature state in two (or perhaps more) domains, the mutual orientation of which possesses a certain degree of freedom. In spite of the above limitations, we find there is a broad range of chain lengths for every set of energetic parameters where the low-temperature state is very specific and resembles the β-barrel structure seen in real proteins. In Figs. 11(A–D) and 12(A–D), the projections of some representative examples of the native structure obtained from computer experiments with $\epsilon_c = \epsilon_g/2$ and $\epsilon_c = \epsilon_g/4$, respectively (model C), are shown for various numbers of structural units (n) in the model molecule.

Some properties that specify the degree of uniqueness of the low-temperature state are compiled in Table III. The collapsed low-temperature structures depicted in Figs. 11 and 12, for $n \leq 200$ are unique in the sense that they satisfy a two-state folding model. That is, fluctuations in the structure that do occur are transparent with respect to the thermodynamic properties. The $n = 50$ chains fold to a structure composed of 4 strands, each of essentially constant length, with very small fluctuations in loop size. Increasing the degree of polymerization tends to decrease the structural uniqueness and the two-state folding picture starts to wash out. However, even up to $n = 200$,

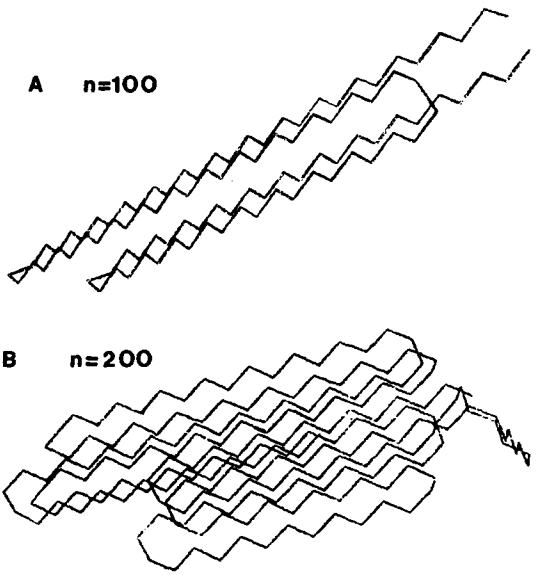


Fig. 12. Representative configuration of the low-temperature states of model C ($\epsilon_a \equiv 0$, $\epsilon_c = \epsilon_g/4$). (A) $n = 100$, (B) $n = 200$, (C–D) $n = 400$ with an illustration of the possibility of the formation of multiple domains in case D.

the bends remain tight and the mean length of a β-strand is almost constant. Thus, the $n = 50$ chains indicate that very highly ordered structures can be formed in the absence of any site-specific interactions. To preserve the degree of structural uniqueness in the sense of the crystallographic location of each of the atoms with increasing n requires additional interactions to be incorporated into the physical model. For example, one way of enhancing the degree of uniqueness might be to specify the location of the bends.

DISCUSSION

Since the simple tetrahedral lattice models of a globular protein described in present work lack site-specific interactions and yet yield well-defined tertiary structures, these MC simulations of equilibrium folding indicate that, contrary to popular wisdom,[36] site-specific interactions are not a necessary condition for the formation of a well-defined, although for larger n not crystallographically unique, tertiary structure. On comparison of models A–C, the existence of some preordering in the denatured state prior to the transition plays an important role in producing well-defined globular structures. In the case of model A, this preordering has to be relatively large when simple, nonbonded nearest neighbor attractions are the only long-range (down the chain) interactions considered, and the presence of substantial secondary structure prior to the transition is necessary to achieve an abrupt collapse to a β-barrellike state. When the chain stiffness is reduced (as required by real β-proteins in the denatured state),[15,30] then the transition is smooth, and the collapsed state resembles a high-density random coil lacking global orientational order of the chain segments. Therefore, there must be another factor(s) that produces the cooperative folding to a well-defined globular structure and

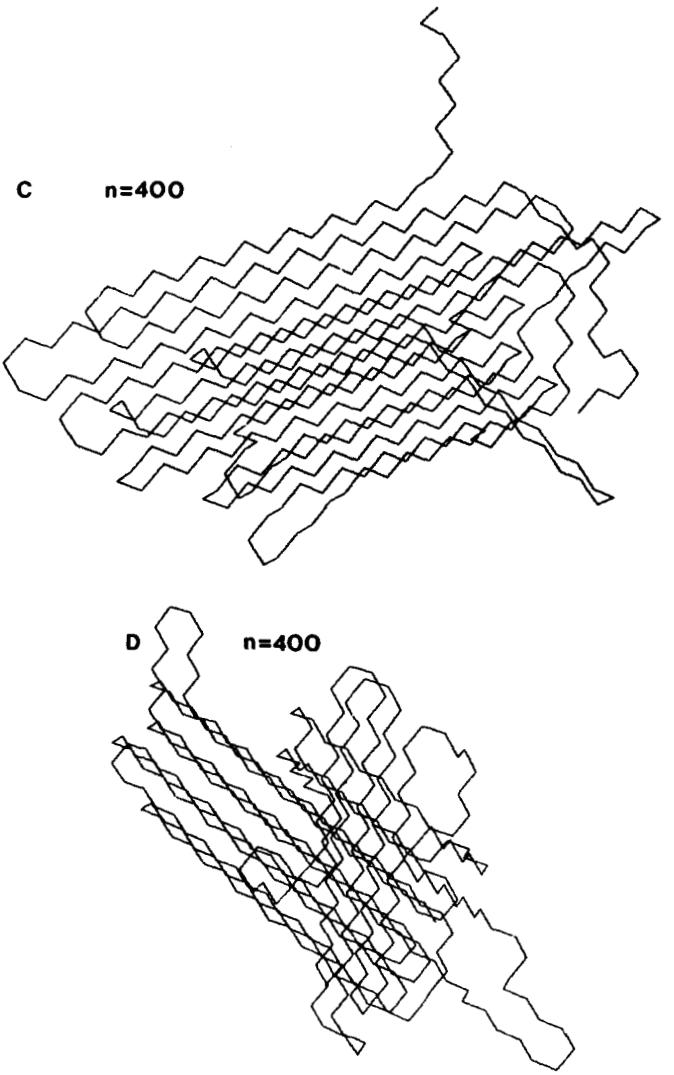


Fig. 12. *(Continued from the previous page.)*

yet preserves the random-coil character of the denatured state. It was demonstrated in model B that a simple model that couples the local nonbonded adjacent conformations leads to a very cooperative, sharp transition to a globular state. However, in the absence of intrinsic local chain stiffness, the global ordering of the low-temperature state is rather marginal. Thus in model C, a small degree of preordering in the denatured state was introduced. The presence of marginal conformational stiffness plus intrachain cooperativity is sufficient to yield a model that reproduces reasonably well the important characteristics of both the high-temperature and low-temperature states of globular proteins. The globular state is well defined, and yet the denatured state is essentially a random coil. These findings are indicative that some small but very important secondary structure in the denatured state may be a

TABLE III
Properties of the Low-Temperature State of Model C

	$n = 50$	$n = 100$	$n = 200$	$n = 400$[a]
$\epsilon_a \equiv 0$, $\epsilon_c = \epsilon_g/2$				
$\langle P_2(\theta)\rangle$[b]	0.37–0.65	0.42–0.60	0.44(±0.03)	0.32(±0.07)
$\langle f_t \rangle$	0.77(±0.07)	0.76(±0.08)	0.74(±0.04)	0.71(±0.09)
$\langle n_t \rangle$	8.9(±2.0)	10.1(±0.07)	11.4(±2.1)	8.1(±1.4)
Number of β-strands	4	4–8	9–14	19–28
$\epsilon_a \equiv 0$, $\epsilon_c = \epsilon_g/4$				
$\langle P_2(\theta)\rangle$	0.73(±0.02)	0.72(±0.03)	0.53(±0.06)	0.44(±0.09)
$\langle f_t \rangle$	0.82(0.01)	0.83(±0.05)	0.80(±0.02)	0.78(±0.01)
$\langle n_t \rangle$	10.8(±0.9)	17.8(±5.4)	12.3(±1.3)	11.9(±2.3)
Number of β-strands	4	4–6	8–12	18–24

[a] Two possible orientations of bundles.
[b] $\langle P_2(\theta)\rangle = 1/2[3 < \cos^2(\mu_i, \mu_j) - 1]$, where $\mu_i = l_i + l_{i+1}$.

necessary condition for the uniqueness of the tertiary structure. Finally, provided that the chain length is moderate, the folding transition is consistent with a two-state model and a β-barrellike native structure emerges. On increasing the chain length, just as in the case of real proteins,[15] the formation of multiple domains is observed.

At this point it is instructive to compare the results obtained here with experiments on homopolymeric polyamino (acids).[38] First, these systems are all at infinite dilution; thus we need not worry about structures formed on aggregation. In contrast, homopolyamino acids such as poly(lysine) almost always aggregate on β-sheet formation; the structure at infinite dilution may or may not resemble those studied here, but in any case the aggregated state tells us nothing about what happens when quaternary interactions are turned off. Moreover, a problem with real homopolyamino acids is that highly β-forming residues tend to be rather hydrophobic.[16] Thus, in homopolymeric form they would either precipitate from solution or aggregate. What must be recognized is that the interaction we employed represents an average over both hydrophobic and hydrophilic residues that produces an equivalent potential of mean force between groups, much as average interhelical interactions parameters have been employed to mimic both salt-bridge and hydrophobic interactions in coiled coils.[38] The potential of mean force employed here as a function of temperature ensures that the random-coil state is the equilibrium high-temperature state and the collapsed state favors β-sheet formation. Hence, rather than viewing the model proteins as being literally homopolymeric, they are in reality the equivalent homopolymeric analog of a real globular protein where the site-specific interactions between residues are replaced by the mean interaction between them. Whether or not a real homopolymeric model system can be found that possesses this kind of interaction remains to be established, but the replacement of the site-specific long-range interactions by an average interaction parameter has ample precedent, and in the case of two-chain coiled coils, works rather well.[39]

How the folding equilibrium can be moderated by the specific distribution of flexibility down the chain backbone and eventually by a site-specific interactions will be discussed in future work. It is hoped these "directed

folding" computational experiments will give us a qualitative understanding of the relative importance of various features in the equilibrium globular protein folding process.

The research described here was supported in part by Grant No. DMR 85-20789 of the Polymer Program of the National Science Foundation and by BRSG S07 RR07054-21 awarded by the Biomedical Research Support Grant Program, Division of Research Resources, National Institutes of Health. Acknowledgment is made to the Monsanto Company for a research grant for the purchase of two μVAX-II computers.

References

1. Lifshitz, I. M., Grosberg, A. Y. & Khoklov, A. R. (1978) *Rev. Mod. Phys.* **50**, 683–695.
2. Schellman, J. A. (1958) *C.R. Trav. Lab. Carlsberg Ser. Chem.* **29**, 230–259.
3. Scheraga, H. A. (1960) *J. Phys. Chem.* **64**, 1917–1926.
4. Tanford, C. (1962) *J. Am. Chem. Soc.* **84**, 4240–4247.
5. Brandts, J. & Lumry, R. (1963) *J. Phys. Chem.* **67**, 1484–1494.
6. Wetlaufer, D. B., Malik, S. K., Stoller, L. & Coffin, R. L. (1964) *J. Am. Chem. Soc.* **86**, 508–514.
7. de Gennes, P. G. (1975) *J. Phys. Lett. (Paris)* **36**, 55–57.
8. Post, C. B. & Zimm, B. H. (1979) *Biopolymers* **18**, 1487–1501.
9. Muthukumar, M. (1984) *J. Chem. Phys.* **81**, 6272–6276.
10. Yamakawa, H. L. & Shimada, J. (1985) *J. Chem. Phys.* **83**, 2607–2611.
11. Birshtein, T. M., Skvorcov, A. M. & Sariban, A. (1977) *Macromolecules* **10**, 202–205.
12. Murakami, H., Norisuye, T. & Fujita, H. (1980) *Macromolecules* **13**, 345–352.
13. Jaenicke, R., Ed. (1980) *Protein Folding, Proceedings of the 28th Conference of the German Biochemical Society*, Elsevier/North Holland Biomedical Press, Amsterdam.
14. Ghelis, C. & Yon, S. (1982) *Protein Folding*, Academic Press, New York.
15. Creighton, T. E. (1985) *J. Phys. Chem.* **89**, 2452–2459.
16. Ptitsyn, O. B. & Finkelstein, V. A. (1980) *Q. Rev. Biophys.* **13**, 339–386.
17. McCammon, J. A. (1984) *Rep. Prog. Phys.* **47**, 1–46.
18. Karplus, M. & McCammon, J. A. (1981) *CRC Crit. Rev. Biochem.* **9**, 293–315.
19. van Gunsteren, W. F. & Berendsen, H. J. C. (1982) *Biochem. Soc. Trans.* **10**, 301–309.
20. Go, N., Abe, H., Mizuno, H. & Taketomi, H. (1980) in *Protein Folding, Proceedings of the 28th Conference of the German Biochemical Society*, Jaenicke, R., ed., Elsevier/North Holland Biomedical Press, Amsterdam, p. 167.
21. Krigbaum, W. R. & Lin, S. F. (1982) *Macromolecules* **15**, 1135–1145.
22. Levitt, M. (1982) *Ann. Rev. Biophys. Bioeng.* **11**, 251–271.
23. Miyazawa, S. & Jernigan, R. L. (1982) *Biopolymers* **21**, 1333–1363.
24. Segawa, S. & Kawai, T. (1986) *Biopolymers* **25**, 1815–1835.
25. Kremer, K., Baumgartner, A. & Binder, K. (1981) *J. Phys. A* **15**, 2879–2883.
26. Binder, K., Ed. (1984) *Application of the Monte Carlo Method in Statistical Physics*, Springer, Berlin, chap. 5.
27. Wall, F. T. & Mandel, F. (1975) *J. Chem. Phys.* **63**, 4592–4595.
28. Kolinski, A., Skolnick, J. & Yaris, R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7267–7271.
29. Kolinski, A., Skolnick, J. & Yaris, R. (1986) *J. Chem. Phys.* **85**, 3585–3597.
30. Tanford, C. (1968) *Adv. Protein Chem.* **23**, 121–282.
31. Flory, P. J. (1969) *Statistical Mechanics of Chain Molecules*, Wiley, New York, chap. 4.
32. Bruns, W. (1984) *Macromolecules* **17**, 2826–2830.
33. McCrackin, F. L., Mazur, J. & Guttman, C. (1973) *Macromolecules* **6**, 859–871.
34. Kolinski, A., Skolnick, J. & Yaris, R. (1987) *Macromolecules*, **20**, 438–440.
35. Privalov, P. L. (1979) *Adv. Protein Chem.* **33**, 167–241.
36. Privalov, P. L. (1982) *Adv. Protein. Chem.* **35**, 1–104.
37. Domb, C. & Fisher, M. E. (1958) *Proc. Cambridge Philos. Soc.* **54**, 48–54.
38. Scaria, P. V., Atreyi, M. & Rao, M. V. (1986) *Biopolymers* **25**, 2349–2358.
39. Skolnick, J. & Holtzer, A. (1985) *Macromolecules* **18**, 1549–1559.

Received November 3, 1986
Accepted January 23, 1987